BOOTSTRAPPING TO OBTAIN CONFIDENCE INTERVALS FOR PARAMETERS IN ORDINARY DIFFERENTIAL EQUATIONS - INFECTIOUS DISEASE MODELS

BU-1502-M **September, 2000**

Martha Contreras
Ruth Zadoks
Heather G. Allore
Ynte H. Schukken

Keywords: nonparametric bootstrap, streptococcus uberis, heteroscedasticity, stochastic simulations, infectious disease modeling.

Abstract: A nonparametric bootstrap is used to construct confidence intervals for the transmission parameters in a nonlinear ordinary differential equation model describing the dynamics of *Streptococcus uberis* infections in a dairy herd. Two levels of susceptibility are modeled and a contagious mode of spread is assumed. *Satisfactory biased-corrected and accelerated* confidence intervals are found for one of the transmission parameters and for the basic reproductive number.

BU1502-M

Bootstrapping to obtain confidence intervals for parameters in ordinary differential equations - infectious disease models

Martha Contreras[1,⋆,⋆⋆], Ruth N. Zadoks[2,⋆⋆], Heather G. Allore[3,⋆⋆], and Ynte H. Schukken[4]

[1]Department of Biometry, Cornell University, 435 Warren Hall, Ithaca, New York 14853, U.S.A., mpc14@cornell.edu.

[2]Department of Farm Animal Health, Faculty of Veterinary Medicine, Utrecht University, Yalelaan 7, 3548 CL Utrecht, The Netherlands, r.n.zadoks@vet.uu.nl.

[3,4]Department of Population Medicine and Diagnostic Sciences, College of Veterinary Medicine, Cornell University, Ithaca, NY 14853, U.S.A., Heather.allore@yale.edu, yhs2@cornell.edu.

SUMMARY.

A nonparametric bootstrap is used to construct confidence intervals for the transmission parameters in a nonlinear ordinary differential equation model describing the dynamics of *Streptococcus uberis* infections in a dairy herd. Two levels of susceptibility are modeled and a contagious mode of spread is assumed. Satisfactory *biased-corrected and accelerated* confidence intervals are found for one of the transmission parameters and for the basic reproductive number.

KEY WORDS: Nonparametric bootstrap, Streptococcus uberis, heteroscedasticity, stochastic simulations, basic reproductive number, transmission parameter, infectious disease modeling.

⋆ Corresponding author and ⋆⋆ denotes that authors contributed equally.

1. Introduction

Streptococcus uberis (Strep. uberis) is a widely occurring causative agent of mastitis in modern dairy herds, leading to loss in milk quality and farm profitability (Barkema et al., 1998, Hogan et al., 1989, McDougall, 1998). In previous work, an outbreak of Strep. uberis mastitis in a dairy herd has been analyzed (Zadoks et al., 2000). This outbreak was described in mathematical terms by a linear model with log-link and Poisson error to predict the number of new infections. In that work, it was found that a Reed-Frost model fitted the observed data better than a Greenwood model. Hence, a contagious mode of spread of the pathogen was assumed (Becker, 1989).

Often times the dynamics of infectious diseases in populations are described by variations of the basic Susceptible - Infectious - Recovered-and-immune (SIR) compartmental model (Anderson and May, 1991). Compartments are mutually exclusive and each individual in the population is in one and only one compartment at any point in time. Transitions between compartments are usually of biological interest. To model different levels of susceptibility in the study population described in this paper, non-infected individuals were assigned to one of two susceptible compartments: "uninfected" for individuals that had not experienced infection before, and "recovered" for individuals that had experienced prior Strep. uberis infection. In contrast to the traditional interpretation of recovered individuals as immune individuals, no complete immunity was assumed after recovery from infection. This is in agreement with field observations from this study and with results from experimental studies (Finch et al., 1997). The rate of becoming infectious depends on composition of the population and on transmission parameters, the probability per unit of time that an infectious quarter will infect a non-infected quarter. Estimation of transmission parameters is then the key interest in these compartmental models.

Typically, transmission parameters are either estimated using generalized linear models (GLM) such as Poisson regression models, or using Ordinary Differential Equations (ODE's). There are important differences between these two methods (Ding an Wu, 2000). In GLM's the rate of new infection is modeled as a linear function of covariates, where these covariates usually include the number of susceptible individuals and the number of infectious individuals just prior to the number of new infections (Haber, Longini, and Cotsonis, 1988, Rampey, 1992). These models essentially use incidence type data, estimate one parameter at a time and provide estimates of variability such that transmission parameters can be evaluated using statistical tests. In contrast, solutions for ODE's provide estimators for multiple parameters simultaneously, but because they are nonlinear in the model parameters they need to be numerically optimized (Brauer and Castillo-Chavez, 2001). Moreover, ODE's are usually based on prevalence type data (changes in compartment size) but do not readily provide estimates of variability for parameters, hence no statistical testing can be performed (Anderson and May, 1991, Nokes and Anderson, 1988).

We will see that it is possible to conduct inference on the model parameters in an ODE via the bootstrap. However, to successfully employ the bootstrap, it is necessary to model the error structure of the data. Because of data limitations, standard tests for heteroscedasticity could not be employed to test the model assumption that the error depended on the mean (Carroll and Ruppert, 1988). Hence, in this paper, we conducted a stochastic simulation study to evaluate the assumption of multiplicative error structure or heteroscedasticity by introducing error around the estimates of parameters. Then, we used the nonparametric bootstrap to construct biased-corrected and accelerated (BCa) confidence intervals of the estimates for the transmission parameters derived from ODE's that characterized the observed outbreak of the infectious

disease, Strep. uberis mastitis in dairy herds. In Section 2 we describe the data and the procedures that we used, in Section 3 the results are presented, and in Section 4 we provide a discussion of our proposed procedure.

2. Data and Modeling Methods

2.1 Data

The data that were used in the models in this paper originate from a longitudinal observational study in a Dutch dairy herd (Zadoks et al., 2000). The herd under study consisted of approximately 100 cows. Information on infection status was based on bacteriological culture from milk samples (Barkema et al., 1998, Harmon et al., 1990). Milk samples were collected at the level of the udder quarter at 3-week intervals. Udder quarters within a cow were treated as independent units, because they are anatomically and physiologically more or less separated, and usually differ in infection status (Sutra and Poutrel, 1994). Also, cross-infections between udder quarters within a cow are not more likely to occur than cross-infections between quarters of different cows (Baxter et al., 1992). Hence, there were approximately 400 individuals in the population at any time during the study. In addition to data on infection status, data on entry of individuals into the herd and exit of individuals from the herd were collected.

Based on current infection status and infection history, all quarters were classified as being infected with Strep. uberis (I), uninfected with no history of Strep. uberis (U), or recovered from infection with Strep. uberis (R). During the 78-week observation period (June 1997 through December 1998), 12,100 quarter milk samples were collected, and 54 infections with Strep. uberis were observed. New infections were observed in

uninfected quarters, as well as in recovered quarters (Figure 1). I_U is used to indicate that a new or existing infection occurred in a previously uninfected quarter. I_R is used to indicate that a new or existing infection occurred in a quarter that had recovered from prior infection. The majority of new infections, i.e. 32 infections, occurred in a limited time period, covering seven 3-week intervals or seven time steps (labeled as times 0,1,2,...,6 on the x-axis of Figure 1). This period constitutes the Strep. uberis outbreak that is described in this paper.

Insert Figure 1 about here

Also depicted in Figure 1 is the size of the infected compartments, I_U and I_R. Compartment size is expressed in units that represent 21 infected quarter-days. This unit is equivalent to one infected quarter that is present for the duration of the 3-week interval or the time step. In the data, quarter-days may have been contributed by more than one quarter, i.e. one quarter may have been infected for 7 days, and a different quarter may have been infected for 14 days in a specified interval. Together, they contribute 21 infected quarter-days to the size of the infected compartment for that time step.

2.2 Ordinary Differential Equation Model

The dynamics of the infection in the population are shown in Figure 2. With respect to the infection of interest, the compartments uninfected (U), infected (I_U or $I_{R)}$ and recovered (R) are distinguished. At each point in time, each individual udder quarter belongs in one compartment. Compartments in the population are mutually exclusive. Transitions between compartments may occur and are indicated by arrows in the model. Uninfected quarters may become infected (transition from U to I_U), infected

quarters may cure (transition from I_U to R), recovered quarters may become reinfected (transition from R to I_R) and cure again (transition from I_R to R). All individuals entering the milking herd are assumed to be non-infected. This is in agreement with the standard assumption that no infected individuals enter the population (Anderson and May, 1991). Therefore, influx into the system occurs only into compartments U and R. Outflow may occur from all compartments. Flow rate is shown for each transition, influx or outflow process. Symbols are explained below.

Insert Figure 2 about here

Assuming independence of quarters within a cow, homogeneous mixing, a constant population size, and constant rates of flow, the compartmental model is described by the following set of differential equations:

$$\begin{aligned}
\frac{d}{dt}U(t) &= (1-q)\mu N + (1-q)\alpha(I_U + I_R) - \beta_U U(\frac{I_U + I_R}{N}) - \mu U \\
\frac{d}{dt}I_U(t) &= \beta_U U(\frac{I_U + I_R}{N}) - (\delta + \gamma + \mu + \alpha)I_U \\
\frac{d}{dt}I_R(t) &= \beta_R R(\frac{I_U + I_R}{N}) - (\delta + \gamma + \mu + \alpha)I_R \\
\frac{d}{dt}R(t) &= q\mu N + q\alpha(I_U + I_R) - \beta_R R(\frac{I_U + I_R}{N}) + (\delta + \gamma)(I_U + I_R) - \mu R,
\end{aligned} \tag{1}$$

where N is the total population size ($U + I_U + I_R + R$), q is the proportion of entries into the system that enter into R, α is the infection associated exit rate, β_U is the transmission parameter for new infections in U, β_R is the transmission parameter for new infections in R, δ is the spontaneous cure rate, γ is the treatment induced cure rate, and lastly μ is the exit rate for exits not associated with infection.

However, we are interested in the behavior of (1) at equilibrium or at steady state. In particular, two types of equilibrium may exist for this system: the *disease free*

equilibrium state or $I = 0$ or the *endemic equilibrium state* or I equal to some nonzero constant (Brauer and Castillo-Chavez, 2001). From a farmer's perspective, the disease free equilibrium is of most interest, as a farmer will strive to control or eradicate Strep. uberis mastitis.

At equilibrium $\frac{dU}{dt} = \frac{dI_U}{dt} = \frac{dI_R}{dt} = \frac{dR}{dt} = 0$, and at the disease free equilibrium the size of the compartments (U, I_U, I_R, R) are $((1-q)N, 0, 0, qN)$, respectively. Because no individuals left the herd solely because of infection, α was considered to be zero. This is in agreement with standard model assumptions (Anderson and May, 1991). Then, for small deviations from the disease free equilibrium, changes in size of the infected compartments are described by the linearization of (1), which leads to

$$\begin{pmatrix} \frac{d}{dt}I_U(t) \\ \frac{d}{dt}I_R(t) \end{pmatrix} = \begin{pmatrix} (1-q)\beta_U - (\delta+\gamma+\mu) & (1-q)\beta_U \\ q\beta_R & q\beta_R - (\delta+\gamma+\mu) \end{pmatrix} \begin{pmatrix} I_U(t) \\ I_R(t) \end{pmatrix}$$

$$:= A\mathbf{I}(t), \tag{2}$$

where A is the 2×2 compartmental matrix whose entries consist of the parameters of the system. System (2) is now a linear ordinary differential equation with initial conditions $(I_U(0), I_R(0)) = (1, 0)$ which correspond to the beginning of the outbreak with 1 infected quarter in the I_U compartment and none in the I_R compartment (see Figure 1). Thus, the solution can be shown to be (see Borrelli and Coleman, 1987)

$$\mathbf{I}(t, \theta, \boldsymbol{\beta}) = e^{A(\theta, \boldsymbol{\beta})t} \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \tag{3}$$

where $(\theta, \boldsymbol{\beta}) = ((q, \delta, \gamma, \mu), (\beta_U, \beta_R))$, respectively, and where θ represents parameters that were assumed to be estimated error-free from the data, and $\boldsymbol{\beta}$ represents the

transmission parameters to be estimated from the ODE model. Lastly, $e^{A(\theta,\beta)t}$ is the *matrix exponential* (Borrelli and Coleman, 1987).

The disease free equilibrium is the desired state that a farmer strives to maintain through adequate herd management. However, cases of disease, or perturbations of the disease free equilibrium, occur. The stability or long term behavior of (2) at the disease free state is then governed by the eigenvalues of the matrix A. More precisely, it is known that a linear ODE is *stable* when $trace(A) < 0$ and $det(A) > 0$ (Borrelli and Coleman, 1987). For a stable equilibrium, incidental cases of disease will not lead to disease outbreaks. Of particular interest is the quantity R_0, or the *basic reproductive number* which indicates whether an epidemic will spread (Brauer and Castillo-Chavez, 2001), and arises when considering the conditions for stability. In fact, the determinant of A yields for our model that the epidemic grows when

$$R_0 := \frac{q\beta_R + (1-q)\beta_U}{\delta + \gamma + \mu} > 1, \tag{4}$$

where the value of R_0 can be interpreted as the number of new infectives that can occur given a single infective.

Typically the parameters of (3) are not known and must be estimated from the data. However, due to error in the data, a measure of the uncertainty in the estimated parameters is needed or a confidence interval needs to be provided. One way to obtain confidence intervals for parameters is via the bootstrap. However, to successfully accomplish this, it is necessary to model the error structure (Carroll and Ruppert, 1988). In the next section, we argue that the proper error structure is multiplicative or that it depends on the mean function. However, we first note that due to data limitations, we separately estimated the parameter θ; that is, values for $(\delta + \gamma)$, or the total cure rate,

and μ were calculated from the data as the number of events, divided by the number of quarter-days at risk for the event. The value for q was estimated from the number of individuals entering the system into compartment R divided by the total number of individuals entering the system. In particular, we let $q = .01$, $(\delta + \gamma) = .169$, and $\mu = .0752$ for the remaining part of our analysis.

2.3 Modeling Error Structure via Stochastic Simulation

Part of the difficulty in properly modeling the error structure of the Strep. uberis data is the small sample size. Nonetheless, one way to address this is to simulate several data sets within a neighborhood of the *nonlinear least squares* (NLS) estimated parameter value, $\hat{\beta}_{NLS}$. Thus, error only enters in the simulated data through the perturbations of the model parameters around the estimated value. This approach is a way of obtaining replicate data sets within the same herd taken at the same time. More precisely, our analysis consisted of the following three steps.

1. Using the previous obtained estimates of θ, we solved

$$\min_{\beta \geq 0} \sum_{i=1}^{m=7} (\mathbf{Y}_i - \mathbf{I}(t_i, \beta))^2 := \hat{\beta}_{NLS},$$

   where β is a 2×1 vector consisting of the transmission parameters, β_U and β_R, which are, as discussed previously, the only parameters that we assumed unknown at this stage of the estimation process, and where $\mathbf{Y}_i$ is a 2×1 vector, per i, consisting of the data in both the I_R and I_U compartments.

2. We then used this estimated value of β to simulate replicate data sets around perturbations of $\hat{\beta}_{NLS}$ according to a *log-normal distribution*; that is, β is *i.i.d.*

$LN(\hat{\beta}_{NLS},\ \sigma^2_{\beta}\mathbf{I}_{2\times 2})$, where the value of $\sigma^2_{\beta} = 1.1$ and error distribution were chosen because they yielded satisfactory data set replicates. That is, we sequentially tested standard deviations and distributions for the parameters until a dataset of 1000 replicates evenly covered the observations of the one observed outbreak.

3. Lastly, we checked the residual fits of the simulated data against their predicted value. Evidence that the variance could depend on mean was noted by the "fanning out" effect discussed in Carroll and Ruppert, 1988 (see Figure 3).

Insert Figure 3 about here

2.4 Bootstrap to Obtain Confidence Intervals

Given the analysis of section 2.3, we can conclude that the modeling assumption of error which depends on the mean is plausible, that is,

$$\mathbf{Y}_i = \mathbf{I}(t_i, \beta) + \epsilon_i \mathbf{I}(t_i, \beta), \quad i = 1, ..., 7, \tag{5}$$

where ϵ_i is normally distributed with mean 0 and constant variance.

To obtain confidence intervals for β and subsequently the basic reproductive number, R_0, we non-parametrically bootstrapped the error, ϵ_i, in (5) upon obtaining an initial estimate of β via *log-of-the-data, log-of-the-model* or the *Transform-Both-Sides* (TBS) methods discussed in Carroll and Ruppert (1988). That is,

$$\epsilon_i \underline{1} = \frac{\mathbf{Y}_i - \mathbf{I}(t_i, \hat{\beta}_{TBS})}{\mathbf{I}(t_i, \hat{\beta}_{TBS})}, \tag{6}$$

are now independently and identically distributed $N(0, \sigma^2)$ random variables, where

$\underline{1}$ is the 2×1 vector of 1's per i, and where $\hat{\beta}_{TBS}$ is the TBS fitted β value. In practice, however, we found more numerically satisfactory results by considering the log-of-the-data, log-of-the-model analogue of (6),

$$\epsilon_i \underline{1} = \frac{\log(\mathbf{Y}_i + \kappa\underline{1}) - \log\left(\mathbf{I}(t_i, \hat{\beta}_{TBS}) + \kappa\underline{1}\right)}{\log\left(\mathbf{I}(t_i, \hat{\beta}_{TBS}) + \kappa\underline{1}\right)}, \tag{7}$$

which are now re-scaled residuals but still independent and identically normally distributed with mean 0 and constant variance, and where we chose $\kappa = 1$ (since this value did not change the inherited structure of the data in the I_U or the I_R compartments, see Figure 4).

We then obtained 10,000 sample data sets based on the residuals in (7) and fitted the log of these data to the log of the model. Lastly, we constructed 95% confidence intervals for the parameters via the *BCa* method. For further details on this bootstrap approach, the reader is referred to Davison and Hinkley (1997) or Efron and Tibshirani (1993). The numerical estimates were obtained using Matlab's *lsqnonlin* optimization routine for nonlinear constrained least squares problems, and the model was encoded using Matlab's built in matrix exponential function, *expm*.

3. Results

In Table 1 we report the results of the bootstrapped estimates. The values of *Bootstrap(mean), TBS, Lower Bound, and Upper Bound* correspond to the mean of the bootstrapped estimates, the original unbootstrapped estimate obtained via TBS methods, the lower bound on the BCa confidence interval, and the upper bound on the confidence interval, respectively.

Insert Table 1 about here

Figure 4 shows the fitted curves to the bootstrapped data and the original unbootstrapped data for both the I_U and the I_R compartments.

Insert Figure 4 about here

Figure 5 has the histograms of the bootstrapped estimates of the transmission parameters, β_U and β_R, and that of the basic-reproductive number R_0.

Insert Figure 5 about here

The confidence interval for R_0 in (4) was computed assuming an a priori error free estimate of q, δ, γ, and μ, as previously discussed, and then using the lower and upper bound estimates of β to compute the lower and upper bound for the estimate of R_0, respectively.

4. Discussion

We believe that the method that we used to obtain an estimate of variability around parameter estimates from ODE's has potentially important applications. Mathematical modeling of infectious diseases using ODE's is frequently practiced to understand the behavior of infectious diseases in populations (Anderson and May, 1991). One of the key parameters in such models in the basic reproduction number, R_0. This parameter indicates that an infection is capable of causing outbreaks when it is larger than 1, whereas values below 1 indicate that an infectious disease agent will not survive in the population. Hence inference about the value of this parameter is of interest. To be successful, control strategies need to reduce the basic reproduction number to a level that does not include values above 1 in the confidence interval (Ferguson et al., 2000).

Using the proposed bootstrapping procedure may provide estimates of variability for the key parameters in mathematical models of disease transmission. Our analysis indicates that given our assumptions R_0 is greater than 1 (see Figure 5). Hence, under current management procedures an outbreak may occur and it may be recommended that preventive measures to control Strep. uberis are implemented.

Point estimates and BCa confidence intervals indicate a higher value for β_R than for β_U. This suggests that recovered individuals are not immune, but, rather more susceptible to infection than individuals that have not experienced infection before. However, the histogram for β_R covers a wide range of bootstrapped estimates, including all values obtained for β_U. Because of that and because of the scarce amount of data available for compartment I_R, we hesitate to conclude that recovered individuals are always more susceptible than other individuals. In fact, considering the bootstrapped estimates for β_R (see Figure 5), it is clear that β_R is zero a significant number of times. This may be a result of the method (log-of-data, log-of-model with $\kappa = 1$) or of the scarcity of data for compartment I_R. From a methodological perspective, we note that using the nonparametric bootstrap with a value of $\kappa > 1$ (e.g. $\kappa = 1.01$) in (7) and then repeating the analysis of Section 3, did yield much smoother histograms for the transmission parameter estimates, but this was at the cost of introducing excessive curvature for the initial values of I_R which, given the actual data observed, we felt did not preserve the structure of the data in the I_R compartment. Further work would be required to better understand the properties of the proposed method.

ACKNOWLEDGMENTS

The corresponding author thanks Charles E. McCulloch for input in the bootstrapping procedure. This work was supported by NSF grant DMS-9710081, Hatch Project NYC-151413, by Intervet International BV, and by Animal Health Service Project 601.936(The Netherlands).

REFERENCES

Anderson, R. M., and May, R. M. (1991). *Infectious Diseases of Human: Dynamics and Control.* Oxford: Oxford University Press.

Barkema, H. W., Schukken, Y. H., Lam, T. J. G. M., Beiboer, M. L., Wilmink, H., Benedictus, G., and Brand, A. (1998). Incidence of clinical mastitis in dairy herds grouped in three categories by bulk milk somatic cell count. *J. Dairy Sci.* **81**, 411-419.

Baxter, J. D., Rogers, G. W., Spencer, S. B., and Eberhart, R. J. (1992). The effect of milking machine liner slip on new intramammary infections. *J. Dairy Sci.* **75**, 1015-1018.

Becker, N. G. (1989). *Analysis of infectious disease data.* London, United Kingdom: Chapman and Hall.

Borrelli, R. L., and Coleman, C. S. (1987). *Differential Equations: A Modeling Approach.* New Jersey: Prentice-Hall, Inc.

Brauer, F., and Castillo-Chavez, C. (2001). *Mathematical Models in Population Biology and Epidemiology.* New York: Springer.

Carroll, R. J., and Ruppert, D. (1988). *Transformation and weighting in regression.* New York: Chapman and Hall.

Davison, A. C., and Hinkley, D. V. (1997). *Bootstrap methods and their application.* New York: Cambridge University Press.

Ding, A. A., and Wu, H. (2000). A comparison study of models and fitting procedures for biphasic viral dynamics in HIV-1 infected patients treated with antiviral therapies. *Biometrics.* Mar;**56**(1), 293-300.

Efron, B., and Tibshirani, R. (1993). *An introduction to the bootstrap.* New York: Chapman and Hall.

Ferguson, N. M., Donnelly, C. A., Woolhouse, M. E., Anderson, R. M. (1999). Estimation of the basic reproduction number of BSE: the intensity of transmission in British cattle. *Proc R Soc Lond B Biol Sci.* Jan **7**;266(1414):23-32.

Finch, J. M., Winter, A., Walton, A. W., and Leigh, J. A. (1997). Further studies on the efficacy of a live vaccine against mastitis caused by *Streptococcus uberis. Vaccine.* **15**(10):1138-43.

Haber, M., Longini, I. M. Jr., Cotsonis, G. A. (1988). Models for the statistical analysis of infectious disease data. *Biometrics.* **44**(1), 163-73.

Harmon, R. J., Eberhart, R. J., Jasper, D. E., Langlois, B. E., and Wilson, R. A. (1990). Microbiological procedures for the diagnosis of bovine udder infection. Arlington, VA: *Nat. Mast. Council.*

Hogan, J. S., Smith, K. L., Hoblet, K. H., Schoenberger, P. S., Todhunter, D. A., Hueston, W. D., Pritchard, D. E., Bowman, G. L., Heider, L. E., Brockett, B.

L., and Conrad, H. R. (1989). Field survey of clinical mastitis in low somatic cell count herds. *J. Dairy Sci.* 72:1547.

McDougall, S. (1998). Efficacy of two antibiotic treatments in curing clinical and subclinical mastitis in lactating dairy cows. *NZ Vet. J.* **46**:226-232.

Nokes, D. J., and Anderson, R. M. (1988). The use of mathematical models in the epidemiological study of infectious diseases and in the design of mass immunization programmes. *Epidemiol Infect.* Aug;**101**(1):1-20.

Rampey, A. H. Jr, Longini, I. M. Jr, Haber, M., and Monto, A. S. (1992). A discrete-time model for the statistical analysis of infectious disease incidence data. *Biometrics.* **48**(1), 117-28.

Sutra, L., and Poutrel B. (1994). Virulence factors involved in the pathogenesis of bovine intramammary infections due to *Staphylococcus aureus. J Med Microbiol* Feb;**40**(2):79-89.

Zadoks, R. N., Allore, H. G., Barkema, H. W., Sampimon, O. C., Gröhn, Y. T., and Schukken, Y. H. (2000). Analysis of an outbreak of Streptococcus uberis mastitis. *Submitted for publication.*

Table 1

Model fitting for Strep. uberis data

			Confidence Intervals	
	Bootstrap(mean)	TBS	Lower Bound	Upper Bound
β_U	.6669	.6860	.5827	.7353
β_R	3.7331	6.5402	4.9666	9.7978
R_0	2.8564	3.0489	2.5657	3.382

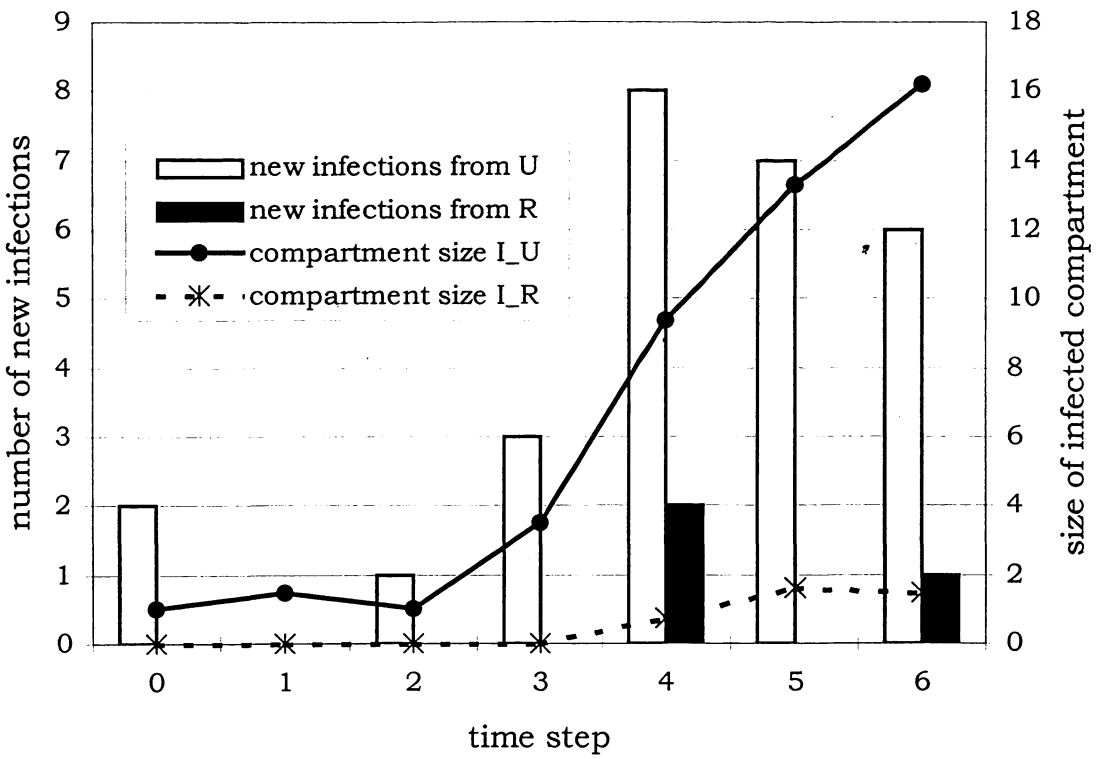

Figure 1: Number of new infections per time step, and size of infected compartments for each time step. I_U represents infections originating in uninfected quarters. I_R represents infections in recovered quarters.

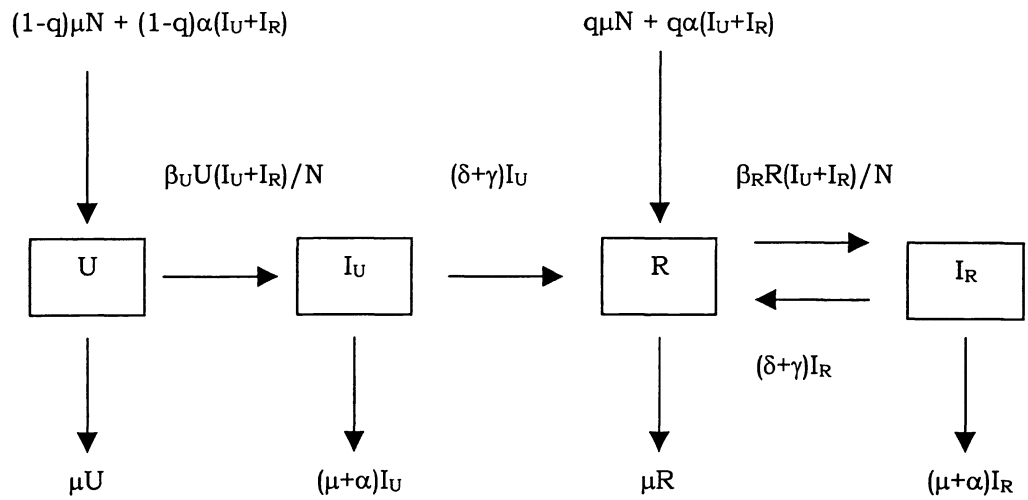

Figure 2: Schematic representation of compartmental model for *Strep. uberis* dynamics in a dairy herd, where N is the total population size ($U+I_U+R+I_R$), q is the proportion of entries into the system that enter into R, α is the infection associated exit rate, β_U is the transmission parameter for new infections in U, β_R is the transmission parameter for new infections in R, δ is the spontaneous cure rate, γ is the treatment induced cure rate, and μ is the exit rate of exits not associated with infection.

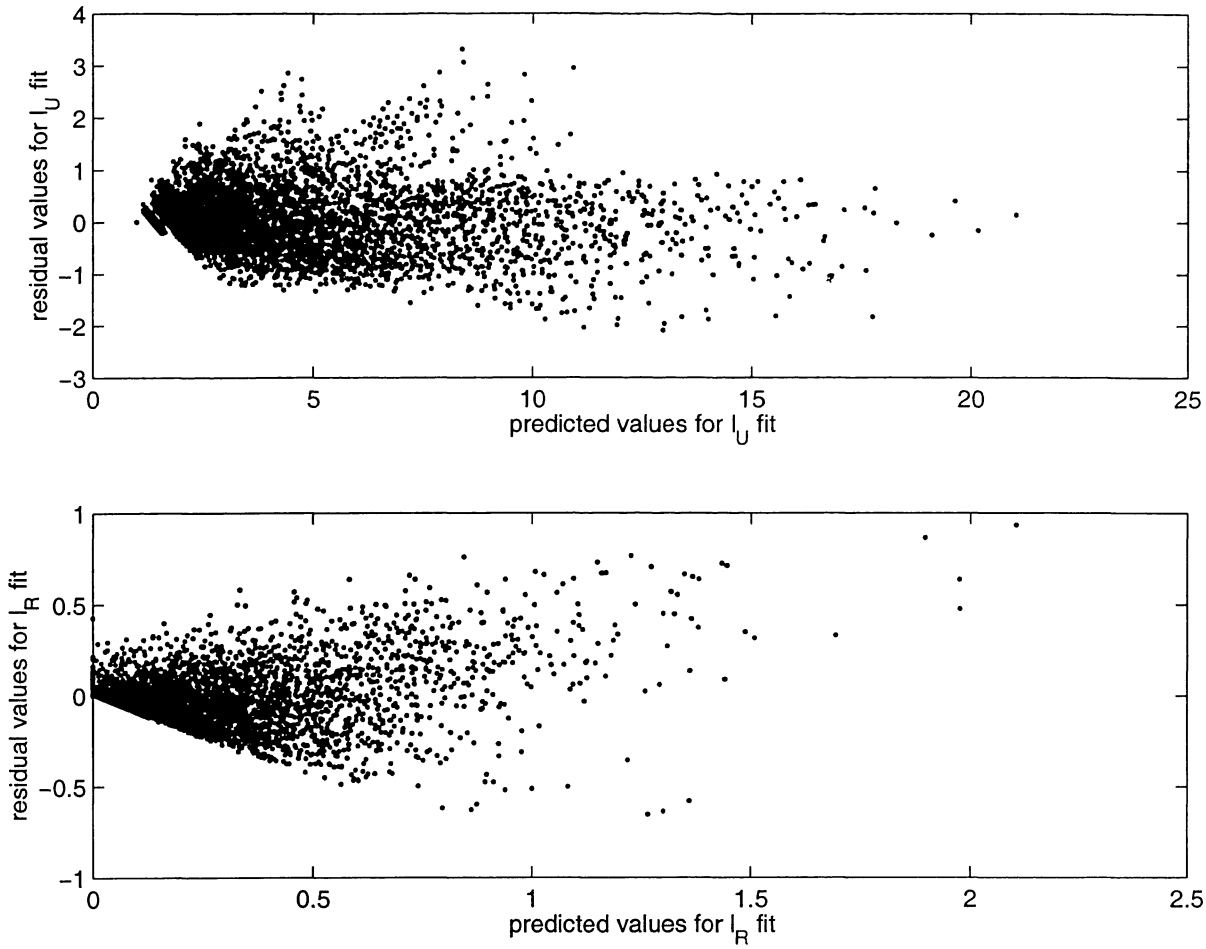

Figure 3: Residual analysis of the stochastically simulated data in both compartments.

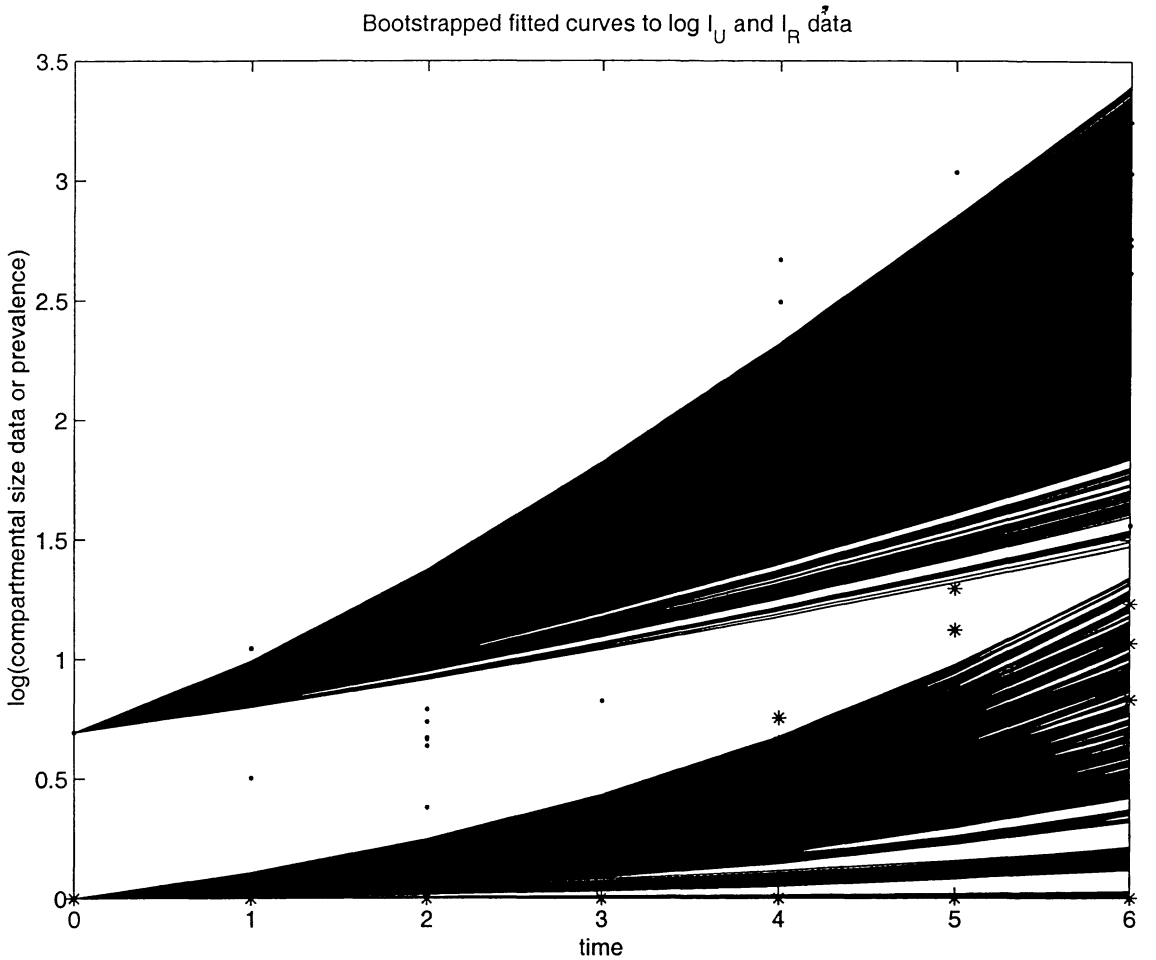

Figure 4: Fitted curves to bootstrapped data in both compartments. '.' and '*' correspond to the bootstrapped data in I_U and I_R compartments, respectively.

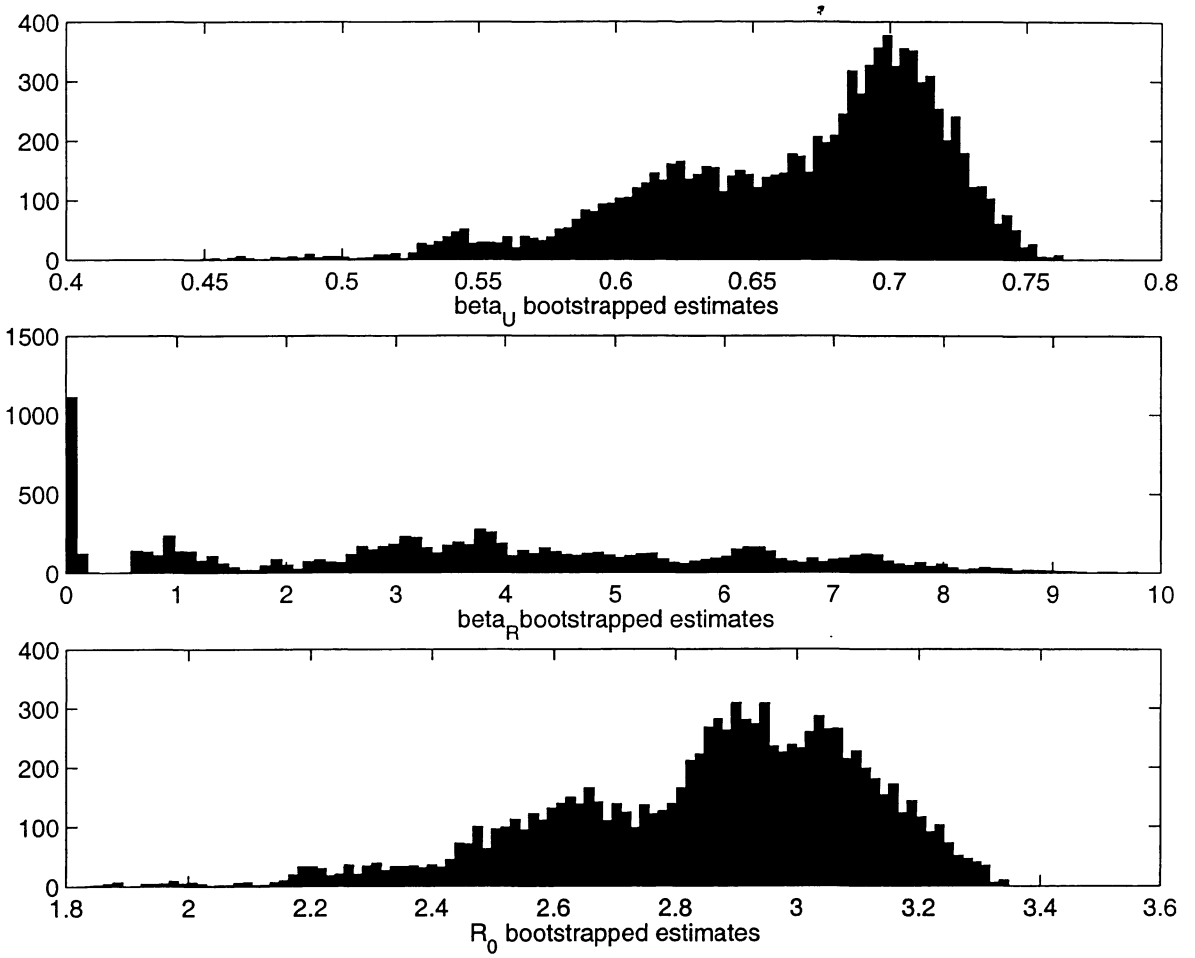

Figure 5: Ten thousand bootstrap estimates for the transmission parameters and the basic reproductive number.